

Optimally weighted realized volatility*

Taro Kanatani^{†‡}

Graduate School of Social Sciences, Hiroshima University
1-2-1 Kagamiyama, Higashi-Hiroshima, 7398525, JAPAN

October 2005

Abstract

In this paper we define a class of estimators for cross-volatility (conditional covariance between two asset returns) by weighted sum of products of two return series. This class nests several estimators and each estimator is characterized by its weight matrix. We derive the MSE-minimizing weight and introduce a feasible example. Our method for measuring cross-volatility is well applicable to nonsynchronous observations.

Keywords: Integrated (cross) volatility; Unevenly sampled observations; Fourier series estimator; Weighted realized volatility

JEL Classification: C14; C32; C63

*The author would like to thank Kimio Morimune, Ryo Okui and seminar participants at Kyoto University, the 2004 Statistics Summer Seminar, the Autumn Meeting of the Japanese Economic Association 2004, the 2004 Annual Meeting of the Japan Statistical Society, and the 2004 Australian Conference of Economists for significant comments and suggestions. This research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology(MEXT), Grant-in-Aid for 21st Century COE Program “Interfaces for Advanced Economic Analysis”.

[†]E-mail: kanatani@hiroshima-u.ac.jp

[‡]URL: <http://www.geocities.jp/kanatanit/>

1 Introduction

Since Black and Scholes (1973) established the theory of option pricing, *volatility*¹ has played an important role not only in the derivatives pricing but also in portfolio selection and risk management. Despite of the assumption of constant volatility in Black and Scholes (1973)², it is widely recognized that volatility changes over time, and other various stylized facts about volatility have been documented (see, e.g., Ghysels, Harvey, and Renault (1996) and Poon and Granger (2003)). These facts have motivated many academic researchers and practitioners to study the dynamics of volatility over the last three decades. Starting with Engle (1982)'s autoregressive heteroskedasticity (ARCH) model, various discrete time models such as Bollerslev (1986)'s generalized ARCH, Nelson (1991)'s exponential ARCH, and stochastic volatility (SV) models have been proposed (see, e.g., Poon and Granger (2003)). On the other hand, volatility is often modeled as a parameterized diffusion coefficient of continuous time diffusion process and then the parameters are estimated via the maximum likelihood methods or general method of moments (see e.g., Lo (1988), Florens-Zmirou (1993), Sueishi (2004)). The link between continuous and discrete time parametric models has been explicitly demonstrated by Drost and Nijman (1993) and Drost and Werker (1996). This paper, however, focuses on nonparametric estimation of volatility process rather than parametric modeling of volatility structure.

In principle, the more data we can use, the more accurate the estimate will be. However, we usually have the technological restriction on the amount of data. Recently this restriction on some kind of financial data has been removing by development of computer power and data recording systems. Those kind of data are called high-frequency data. Such high-frequency data

¹Throughout this paper, we use the term “volatility” to reference both variance (not standard deviation) and covariance.

²Hull and White (1987) modifies Black and Scholes (1973)'s option pricing formula for stochastic volatility.

lend the validity to the method based on quadratic variation formula, that is called as *realized volatility* in the finance and econometrics literature. We concentrate on the *ex post* volatility measuring by these type of methods.

Barndorff-Nielsen and Shephard (2004) derives asymptotic distribution of realized volatility matrix — the sum of outer products of high frequency vectors of returns. Since their purpose is to provide the asymptotic distribution theory, they establish the theory for data observed at equally spaced time intervals. Andersen, Bollerslev, Diebold, and Labys (2003) provide methods of realized volatility incorporated into lower frequency volatility models. For example, Using intradaily observations for the Deutschemark/Dollar and Yen/Dollar spot exchange rates, they find that forecasts from a long memory Gaussian vector autoregression for the logarithmic daily volatilities perform admirably.

While all of the theories mentioned above are built on the evenly sampled observations, Malliavin and Mancino (2002) proposed an estimator base on Fourier series analysis that is well suited for unevenly sampled observations, in other words, for tick-by-tick data without any data manipulation. One of the most important purpose to use tick-by-tick data is to avoid the interpolation bias. Because of the facility of handling, tick-by-tick (transaction) data, which inherently arrive in irregular time intervals, are usually transformed into regularly spaced data through a certain interpolation. However, that interpolation method reduces the number of data and introduces the bias. The bias is serious especially in cases of measuring cross volatility. Hayashi and Yoshida (2005) proposed an unbiased nonsynchronous covariance estimator and studied its asymptotic properties. We generalize all these quadratic-variation-based methods and seek more accurate estimator on the basis of mean squared error.

The layout of this paper is as follows. In Section 2 we define weighted realized volatility as a estimator of integrated cross volatility and show that it nests several estimators such as Fourier series estimator of Malliavin and

Mancino (2002) and realized volatility based on interpolated returns. In Section 3 we derive the MSE-minimizing weight and provide a feasible example of it. Through a Monte Carlo simulation, we examine our theory in Section 4. Section 5 summarizes this paper and overviews future studies.

1.1 Data generating process and observations

We consider n -dimensional logarithmic price $p(t) = (p_1(t), \dots, p_n(t))'$ which follows the stochastic differential equation:

$$dp(t) = \Sigma(t) dz(t), \quad 0 \leq t \leq T$$

where $\Sigma(t)$ is an $n \times n$ matrix $[\sigma_{ij}(t)]_{i,j=1,\dots,n}$, and z is an $n \times 1$ vector of independent standard Brownian motions. We set the drift vector as 0 for the purpose of simplification.³ We define the volatility matrix as

$$\Omega \equiv \Sigma \Sigma',$$

that is to say, cross volatility between i th and j th asset is denoted as the ij element of Ω :

$$\omega_{ij}(t) = \sum_{k=1}^n \sigma_{ik}(t) \sigma_{jk}(t).$$

Each i th asset price is observed at irregular time points $\{t_k^i\}_{k=0}^{N_i}$.⁴ We just impose the assumption on the observation points that the time intervals are small: $\lim_{N_i \rightarrow \infty} \sup_{j \geq 1} (t_j^i - t_{j-1}^i) = 0$. Since we concentrate on the *ex post* cross volatility measuring and do not make any hypothesis on the structure of the underlying probability space Ω , we can construct an auxiliary probability space X where we consider $\Sigma(t)$ as deterministic functions. See Malliavin and Mancino (2002). Throughout this paper, E denotes the expectation on the probability space X .

³This simplification is acceptable not only because it means an efficient market in financial economics, but also because, mathematically, the martingale component swamps the predictable portion over short time intervals.

⁴For the purpose of simplification, we set $t_0^i = 0$ and $t_{N_i}^i = T$.

2 Weighted realized volatility

2.1 Representation

We define the estimator of $\int_0^T \omega_{ij}(t) dt$ as

$$\hat{\omega}_{ij} = \Delta p_i' W \Delta p_j = \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} w_{kl} \Delta p_i(t_k^i) \Delta p_j(t_l^j) \quad (2.1)$$

where

$$\Delta p_i = \begin{pmatrix} p_i(t_1^i) - p_i(t_0^i) \\ \vdots \\ p_i(t_{N_i}^i) - p_i(t_{N_i-1}^i) \end{pmatrix}, W = \begin{pmatrix} w_{11} & \cdots & w_{1N_j} \\ \vdots & \ddots & \vdots \\ w_{N_i1} & \cdots & w_{N_iN_j} \end{pmatrix}.$$

We call (2.1) *weighted realized volatility*. (2.1) nests several estimators of the integrated volatility $\int_0^T \omega_{ij}(t) dt$. For example, if $w_{ij} = 1$ for any k, l ,

$$\begin{aligned} \hat{\omega}_{ij} &= \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} \Delta p_i(t_k^i) \Delta p_j(t_l^j) \\ &= \left\{ \sum_{k=1}^{N_i} \Delta p_i(t_k^i) \right\} \left\{ \sum_{l=1}^{N_j} \Delta p_j(t_l^j) \right\} \\ &= \{p_i(t_{N_i}) - p_i(t_0)\} \{p_j(t_{N_j}) - p_j(t_0)\} \\ &= \{p_i(T) - p_i(0)\} \{p_j(T) - p_j(0)\} \end{aligned} \quad (2.2)$$

which is an unbiased but very noisy estimator of $\int_0^T \omega_{ij}(t) dt$. If the window $[0, T]$ is one day, (2.2) means that we measure daily (cross) volatility by using daily return, in other words, discarding all intradaily data of $\{p_i(t_k^i)\}_{k=1}^{N_i-1}$. In this manner, the weight matrix characterizes the data for measuring volatility. In order to understand this point, we look at another example. In univariate settings, if $W = I_{N_i}$,

$$\hat{\omega}_{ii} = \sum_{k=1}^{N_i} (p_i(t_k^i) - p_i(t_{k-1}^i))^2.$$

Note that this estimator uses all available observations, therefore, is expected to be less noisy. We discuss the multivariate version of this in the subsection 2.4. Throughout the following three subsections, the examples of (2.1) are discussed.

2.2 Interpolation and realized volatility

The raw data which are unevenly spaced, are converted to evenly spaced data in order to apply to the usual discrete time series analysis. Dacorogna, Gençay, Müller, Olsen, and Pictet (2001) introduces some interpolation methods including *linear interpolation* and *previous tick interpolation*.⁵ When constructing M evenly spaced data $\{q(mT/M)\}_{m=0}^M$ from $\{p_i(t_k^i)\}_{k=0}^{N_i}$, those data manipulation is as follows:

$$q_i\left(\frac{mT}{M}\right) = \begin{cases} (1 - \rho_m^i) p_i(*t_m^i) + \rho_m^i p_i(*t_m^i) & \text{linear interpolation} \\ p_i(*t_m^i) & \text{previous-tick interpolation} \end{cases} \quad (2.3)$$

where

$$\begin{aligned} \rho_m^i &= \frac{(mT/M) - *t_m^i}{*t_m^i - *t_m^i}, \\ *t_m^i &= \max \{t_k^i : t_k^i \leq mT/M\}, \\ *t_m^i &= \min \{t_k^i : t_k^i \geq mT/M\}, \end{aligned}$$

and where $\max A$ and $\min A$ denote maximum and minimum elements of A , respectively.

Using evenly spaced data of $\{q_i(mT/M)\}_{m=0}^M$ and $\{q_j(mT/M)\}_{m=0}^M$, the integrated cross volatility $\int_0^T \omega_{ij}(t)dt$ is measured by the following estimator,

$$\hat{\omega}_{ij}(M) = \sum_{m=1}^M \left(q_i\left(\frac{mT}{M}\right) - q_i\left(\frac{(m-1)T}{M}\right) \right) \left(q_j\left(\frac{mT}{M}\right) - q_j\left(\frac{(m-1)T}{M}\right) \right). \quad (2.4)$$

⁵Dacorogna, Gençay, Müller, Olsen, and Pictet (2001) also introduces *next tick interpolation* which is analogous to previous tick interpolation.

In order to distinguish difference on the interpolation procedure, we introduce the notation of $\hat{\omega}_{ij}^L(M)$ and $\hat{\omega}_{ij}^P(M)$ for liner interpolation and previous-tick interpolation, respectively. Barucci and Renò (2002) shows through Monte Carlo simulation that $\hat{\omega}_{ii}^L(M)$ has the downward bias. Kanatani (2004) calculates the theoretical bias. As we use higher and higher frequency data, the bias becomes more profound. Thus, the linear interpolation is not suitable for calculation of realized volatility.

On the other hand, $\hat{\omega}_{ii}^P(M)$ is unbiased. The bias of $\hat{\omega}_{ij}^P(M)$ is

$$\sum_{m=1}^M \int_{t_m^-}^{t_m^+} \omega_{ij}(t) dt \quad (2.5)$$

where

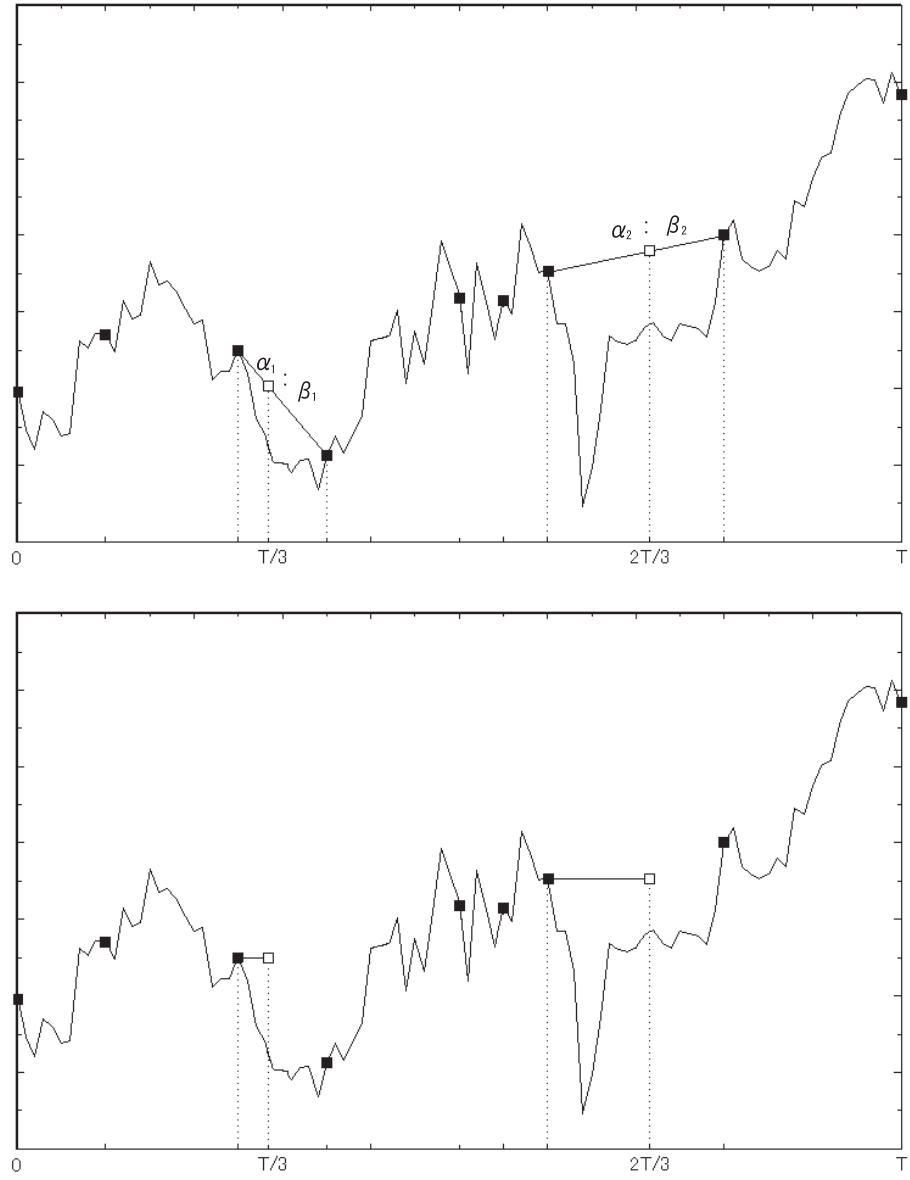
$$\begin{aligned} t_m^- &= \min \{ {}_*t_m^i, {}_*t_m^j \}, \\ t_m^+ &= \max \{ {}_*t_m^i, {}_*t_m^j \}. \end{aligned}$$

Notice that in the case of univariate volatility ($i = j$), for $t_m^- = t_m^+$, the realized volatility through previous tick interpolation is an unbiased estimator.

In order to show that the realized volatility (2.4) can be written by the expression of the weighted realized volatility (2.1), we shall present a simple example.

Example 1 *Let us consider a simple case as shown in Figure 1: $M = 3, N_i = 8$. $\hat{\omega}_{ii}^L(M)$ can be written by the form of weighted realized volatil-*

Figure 1: Linear interpolation and Previous-tick interpolation



Note: Linear interpolation (upper) and Previous-tick interpolation (lower). Black and white squares denote observed and interpolated data respectively.

ity (2.1) with the weight matrix:

$$W = \begin{pmatrix} 1 & 1 & \alpha_1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & \alpha_1 & 0 & 0 & 0 & 0 & 0 \\ \alpha_1 & \alpha_1 & \alpha_1^2 + \beta_1^2 & \beta_1 & \beta_1 & \beta_1 & \beta_1 \alpha_2 & 0 \\ 0 & 0 & \beta_1 & 1 & 1 & 1 & \alpha_2 & 0 \\ 0 & 0 & \beta_1 & 1 & 1 & 1 & \alpha_2 & 0 \\ 0 & 0 & \beta_1 & 1 & 1 & 1 & \alpha_2 & 0 \\ 0 & 0 & \beta_1 \alpha_2 & \alpha_2 & \alpha_2 & \alpha_2 & \alpha_2^2 + \beta_2^2 & \beta_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & \beta_2 & 1 \end{pmatrix}. \quad (2.6)$$

See Appendix A.1 for the detail derivation of (2.6). Since previous tick interpolation is a special case of the linear interpolation for $\alpha_m = 0$ and $\beta_m = 1$, $\hat{\omega}_{ii}^P(M)$ can be written by the form of weighted realized volatility (2.1) with the weight matrix:

$$W = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}. \quad (2.7)$$

2.3 Fourier series estimator of Malliavin and Mancino (2002)

Malliavin and Mancino (2002) proposes a new method for measuring volatility by using Fourier analysis. The method is especially suitable for unevenly sampled observations. We prove that the Fourier series estimator can be written by the form of the weighted realized volatility. In this subsection we normalized the time window $[0, T]$ to $[0, 2\pi]$. Fourier series estimator of

Malliavin and Mancino (2002) for the integrated (cross) volatility $\int_0^{2\pi} \omega_{ij} dt$ is as follows.

$$\hat{\omega}_{ij}^F = \frac{\pi^2}{Q} \sum_{q=1}^Q (a_q(dp_i) a_q(dp_j) + b_q(dp_i) b_q(dp_j)) \quad (2.8)$$

where

$$a_q(dp_i) = \frac{1}{\pi} \int_0^{2\pi} \cos(qt) dp_i(t), \quad (2.9)$$

$$b_q(dp_i) = \frac{1}{\pi} \int_0^{2\pi} \sin(qt) dp_i(t), \quad (2.10)$$

and Q is a large integer. We will compute the Fourier coefficient (2.9) through integration by parts:

$$\begin{aligned} a_q(dp_i) &= \frac{1}{\pi} \int_0^{2\pi} \cos(qt) dp(t) \\ &= \frac{p_i(2\pi) - p_i(0)}{\pi} + \frac{1}{\pi} \int_0^{2\pi} \sin(qt) p_i(t) dt \\ &\approx \frac{p_i(2\pi) - p_i(0)}{\pi} + \frac{1}{\pi} \sum_{j=0}^{N_i-1} [\cos(qt_j^i) - \cos(qt_{j+1}^i)] p_i(t_j^i), \end{aligned}$$

since the piecewise constant is valid under assumption $\lim_{N \rightarrow \infty} \sup_{j \geq 1} (t_j^i - t_{j-1}^i) = 0$. Similarly, we approximate (2.10) by

$$b_q(dp_i) \approx \frac{1}{\pi} \sum_{j=0}^{N-1} [\sin(qt_j^i) - \sin(qt_{j+1}^i)] p_i(t_j^i). \quad (2.11)$$

This approximation of the integrals is proved to be equivalent to setting the weight in (2.1) as follows.

$$w_{kl} = \begin{cases} 1 & \text{if } t_k^i = t_l^j, \\ \frac{\sin \frac{(Q+1)(t_k^i - t_l^j)}{2} \cos \frac{Q(t_k^i - t_l^j)}{2}}{Q \sin \frac{(t_k^i - t_l^j)}{2}} & \text{otherwise.} \end{cases}$$

See the Appendix A.2. In the special case of univariate volatility ($i = j$), as we increase the number of Fourier coefficients ($Q \rightarrow \infty$), the weight matrix converges to identity matrix ($W \rightarrow I_{N_i}$). In the case of cross volatility ($i \neq j$), since transaction is usually nonsynchronous, $t_k^i - t_l^j$ has some width. Therefore, as $K \rightarrow \infty$, $w_{kl} \rightarrow 0$: $\hat{\omega}_{ij}^F \rightarrow 0$. Thus we should not increase the number of Fourier coefficients.

2.4 Raw data realized volatility

Another method for measuring integrated volatility $\int_0^T \omega_{ii} dt$ using unevenly sampled observations $\{p_i(t_k^i)\}_{k=0}^{N_i}$ is

$$\hat{\omega}_{ii}^R = \sum_{i=1}^{N_i} \{\Delta p_i(t_k^i)\}^2$$

As described in Subsection 2.1, this estimator is also written by the form of weighted realized volatility with identity matrix I_{N_i} . Kanatani (2004) provides the relationship between raw data realized volatility and Fourier series estimator:

$$\hat{\omega}_{ii}^F \rightarrow \hat{\omega}_{ii}^R \text{ and } V(\hat{\omega}_{ii}^F) \downarrow V(\hat{\omega}_{ii}^R) \text{ as } Q \rightarrow \infty.$$

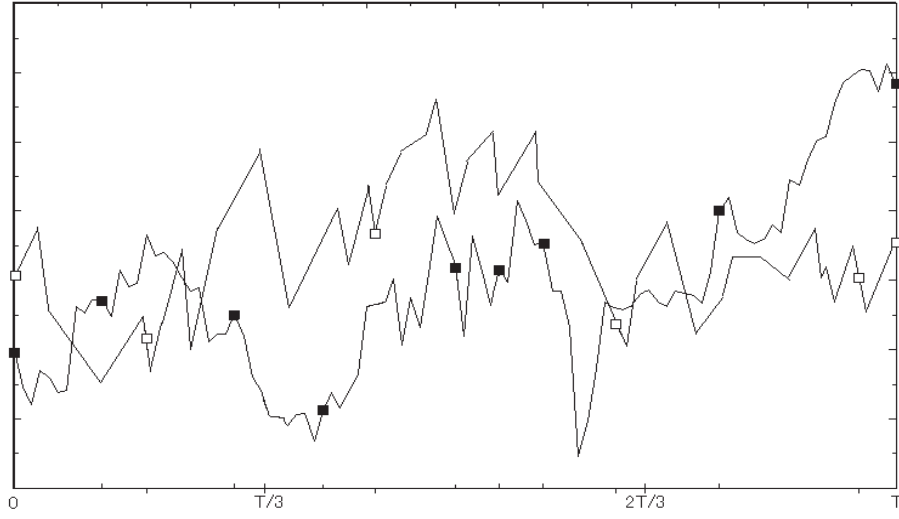
For measuring cross volatility, we extend the method using unevenly sampled observations $\{p_i(t_k^i)\}_{k=0}^{N_i}$ and $\{p_j(t_l^j)\}_{l=0}^{N_j}$:

$$\hat{\omega}_{ij}^R = \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} \Delta p_i(t_k^i) \Delta p_j(t_l^j) I(A) \quad (2.12)$$

where $A = \{(t_k^i, t_{k-1}^i) \cap (t_l^j, t_{l-1}^j) \neq \emptyset\}$ and $I(\cdot)$ denotes indicator function. We refer to (2.12) as *raw data realized (cross) volatility*. (2.12) is expressed by the form of weighted realized volatility with the weights:

$$w_{kl} = \begin{cases} 1 & \text{if } (t_k^i, t_{k-1}^i) \cap (t_l^j, t_{l-1}^j) \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 2: Nonsynchronous observations



Note: Black and white squares denote observed data of 1st and 2nd asset respectively.

Although all estimators of cross volatility mentioned throughout the previous subsections introduce the bias, this simple estimator is constructed to be unbiased. Hayashi and Yoshida (2005) proves its consistency.

Example 2 *Let us consider a bivariate case as shown in Figure 2: $M = 3, N_1 = 8, N_2 = 5$. $\hat{\omega}_{21}^R$ can be written by the form of weighted realized volatility (2.1) with the weight matrix:*

$$W = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (2.13)$$

$\hat{\omega}_{21}^P(M)$ can be written by the form of weighted realized volatility (2.1) with

the weight matrix:

$$W = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}. \quad (2.14)$$

3 Optimal weight

3.1 MSE-minimizing weight

In this subsection, we derive the optimal weight that minimizes the MSE of (2.1):

$$E \left(\hat{\omega}_{ij} - \int_0^T \omega_{ij}(t) dt \right)^2 = bias^2 + V(\hat{\omega}_{ij}).$$

We define the intersection interval as

$$I(k, l) \equiv (t_k^i, t_{k-1}^i) \cap (t_l^j, t_{l-1}^j).$$

We introduce a convenient notation for the element of weight matrix W as follows

$$\begin{cases} w_{kl}^A & \text{if } I(k, l) \neq \emptyset, \\ w_{kl}^{AC} & \text{otherwise.} \end{cases}$$

The bias is given by

$$\begin{aligned} & E \left(\sum_{k=1}^{N_i} \sum_{l=1}^{N_j} w_{kl} \Delta p_i(t_k^i) \Delta p_j(t_l^j) \right) - \int_0^T \omega_{ij}(t) dt \\ &= E \left(\sum w_{kl}^A \int_{t_{k-1}^i}^{t_k^i} dp_i(t) \int_{t_{l-1}^j}^{t_l^j} dp_j(t) \right) - \int_0^T \omega_{ij}(t) dt \\ &= \sum w_{kl}^A \int_{I(k,l)} \omega_{ij} dt - \int_0^T \omega_{ij}(t) dt. \end{aligned} \quad (3.1)$$

Note that if $w_{kl}^A = 1$, the bias is zero. On the other hand, the variance is given by

$$\begin{aligned}
& V(\hat{\omega}_{ij}) \\
&= \sum \left\{ \left(\int_{I(k,l)} \omega_{ij} dt \right)^2 + \left(\int_{t_{k-1}}^{t_k} \omega_{ii} dt \right) \left(\int_{t_{l-1}}^{t_l} \omega_{jj} dt \right) \right\} (w_{kl}^A)^2 \\
&+ \sum \left(\int_{t_{k-1}}^{t_k} \omega_{ii} dt \right) \left(\int_{t_{l-1}}^{t_l} \omega_{jj} dt \right) (w_{kl}^{AC})^2. \tag{3.2}
\end{aligned}$$

See the Appendix A.3. It is obvious that we should set $w_{kl}^{AC} = 0$ in order to minimize the MSE because $\omega_{ii}(t)$ is nonnegative.

For example, compare the identity matrix with (2.14). Although diagonal element of these two are identical, (2.14) has some non-zero off-diagonal elements ($w_{kl}^{AC} \neq 0$). This means that variance of previous-tick realized volatility is larger than raw data realized volatility. As another example, remember the weight matrices of (2.8) and (2.12) in the case of univariate volatility. Both of them have the same diagonal elements $w_{kl}^A = 1$, while (2.8) have non-zero off-diagonal elements w_{kl}^{AC} . Therefore, variance of (2.8) is larger than that of (2.12). As $Q \rightarrow \infty$, w_{kl}^{AC} of (2.8) goes to 0, then these two are almost same. See Kanatani (2004).

In order to minimize the MSE, we set $w_{kl}^{AC} = 0$ and then rewrite the MSE in matrix expression.

$$MSE = w' D w + (x' (w - 1))^2$$

where

$$\begin{aligned}
w &= \begin{pmatrix} w_{11}^A \\ \vdots \\ w_{kl}^A \\ \vdots \\ w_{N_i N_j}^A \end{pmatrix}, D = \begin{pmatrix} v_{11} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & v_{kl} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & v_{N_i N_j} \end{pmatrix}, \\
x &= \begin{pmatrix} \int_{I(1,1)} \omega_{ij} dt \\ \vdots \\ \int_{I(k,l)} \omega_{ij} dt \\ \vdots \\ \int_{I(N_i, N_j)} \omega_{ij} dt \end{pmatrix}, 1 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \\
v_{kl} &= \left(\int_{I(k,l)} \omega_{ij} dt \right)^2 + \left(\int_{t_{k-1}}^{t_k} \omega_{ii} dt \right) \left(\int_{t_{l-1}}^{t_l} \omega_{jj} dt \right).
\end{aligned}$$

Let

$$u_{kl} = \frac{\left(\int_{I(k,l)} \omega_{ij} dt \right)^2}{v_{kl}},$$

then we get the following theorem.

Theorem 3 *The MSE of (2.1) is globally minimized by the following weight:*

$$w_{kl}^A = \frac{\int_0^T \omega_{ij} dt \int_{I(k,l)} \omega_{ij} dt}{v_{kl} \{1 + \sum u_{kl}\}}. \quad (3.3)$$

The bias and variance obtained by using the optimal weight (3.3) are

$$\frac{-\int_0^T \omega_{ij} dt}{1 + \sum u_{kl}} \quad \text{and} \quad (3.4)$$

$$\left(\frac{\int_0^T \omega_{ij} dt}{1 + \sum u_{kl}} \right)^2 \sum u_{kl}, \quad (3.5)$$

respectively. The minimized MSE is

$$\frac{\left(\int_0^T \omega_{ij} dt\right)^2}{1 + \sum u_{kl}}. \quad (3.6)$$

Proof. See Appendix A.4 ■

In order to understand the property of the optimal weight, consider a special case of the individual volatility ($i = j$). Since

$$v_{kk} = 2 \left(\int_{t_{k-1}}^{t_k} \omega_{ij} dt \right)^2 \quad \text{and} \quad u_{kk} = \frac{1}{2},$$

W is an $N_i \times N_i$ diagonal matrix that has diagonal elements

$$w_{kk}^A = \frac{1}{N_i + 2} \frac{\int_0^T \omega_{ii} dt}{\int_{t_{k-1}}^{t_k} \omega_{ii} dt}.$$

This weight increases (decreases) when $\int_{t_{k-1}}^{t_k} \omega_{ii} dt$ decreases (increases). This fact implies that larger (smaller) weights are assigned in densely (coarsely) sampled periods and that smaller (larger) weights are assigned in volatile (less volatile) periods. The bias and variance are

$$\begin{aligned} & \frac{-2}{N_i + 2} \int_0^T \omega_{ii}(t) dt \quad \text{and} \\ & \frac{2N_i}{(N_i + 2)^2} \left(\int_0^T \omega_{ii}(t) dt \right)^2, \end{aligned}$$

respectively. The estimator is not unbiased, however, the bias shrinks at order $1/N_i$. The variance also shrinks at order $1/N_i$ in similar fashion to the variance of *realized variance* of Barndorff-Nielsen and Shephard (2004)⁶.

⁶Barndorff-Nielsen and Shephard (2004) studies properties of sum of squared returns in the case of evenly sampled observations. They refer to the sum of square returns as realized variance and to square root of it as realized volatility.

3.2 An estimator of nuisance parameters

To construct the optimal weight of Theorem 3, we must estimate $\int_{I(k,l)} \omega_{ij} dt$. We call it *piecewise integrated volatilities* (PWIV). It is essentially difficult to estimate them. We give an example of unbiased estimators: we use

$$\Delta p_i(t_k^i) \Delta p_j(t_l^j) \text{ and } \{\Delta p_i(t_k^i)\}^2,$$

as estimators of $\int_{I(k,l)} \omega_{ij} dt$ and $\int_{t_{k-1}}^{t_k} \omega_{ii} dt$ respectively. We also need an estimate of $\int_0^T \omega_{ij}(t) dt$, in Monte Carlo study of next section, we use (2.12). By using these estimators to construct the weights, the weighted realized volatility (2.1) is equivalent to

$$\hat{\omega}_{ij}^N = \frac{N_{ij} \hat{\omega}_{ij}^R}{N_{ij} + 2} \quad (3.7)$$

where $N_{ij} = N_i + N_j - \sum I(\{t_k^i = t_l^j\})$. We refer to (3.7) as naively weighted realized volatility. Although there is little difference between $\hat{\omega}_{ij}^N$ and $\hat{\omega}_{ij}^R$ when N_{ij} is large, we find that $\hat{\omega}_{ij}^N$ slightly improves the MSE compared with $\hat{\omega}_{ij}^R$ in the Monte Carlo study of next section.

4 Monte Carlo study

We examine the above theory through a Monte Carlo study. Without loss of generality, we set the number of assets as two. We follow the Monte Carlo design of Barucci and Renò (2002) with some modification for multivariate setting: we generate proxy for continuous observations by discretizing following stochastic differential equations with a time step of one second,

$$\begin{pmatrix} dp_1(t) \\ dp_2(t) \end{pmatrix} = \begin{pmatrix} \sigma_{11}(t) & \sigma_{12}(t) \\ \sigma_{21}(t) & \sigma_{22}(t) \end{pmatrix} \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix}, \quad 0 \leq t \leq T$$

$$d\sigma_{ij}(t) = \kappa_{ij}(\theta_{ij} - \sigma_{ij}(t))dt + \gamma_{ij}dW_{ij}(t), \quad i, j = 1, 2.$$

where $\kappa_{ij} = 0.01$, $\theta_{ij} = 0.01$, and $\gamma_{ij} = 0.001$ for any i, j and $T = 60 \times 60 \times 24$ seconds. Time differences are drawn from an exponential distribution with mean 45 seconds for p_1 and 60 seconds for p_2 .⁷

$$F(t_k^i - t_{k-1}^i) = 1 - \exp\{-\lambda_i(t_k^i - t_{k-1}^i)\}, \quad i = 1, 2$$

where $F(\cdot)$ denotes a cumulative distribution function, $\lambda_1 = 1/45$ and $\lambda_2 = 1/60$.

We compared the performances of previous tick interpolation realized volatility $\hat{\omega}_{ij}^P(M)$, Fourier series estimator $\hat{\omega}_{ij}^F$, raw data realized volatility $\hat{\omega}_{ij}^R$, and naively weighted realized volatility $\hat{\omega}_{ij}^N$. We also observed the performance of the estimator using the optimal weight. In calculations of previous tick interpolation realized volatility $\hat{\omega}_{ij}^P(M)$, we set $M = 144, 288$, and 720 , corresponding to so-called daily realized volatility based on 10-min, 5-min and 2-min returns. In calculations of Fourier series estimator $\hat{\omega}_{ij}^F$, we set $Q = 10, 25, 50, 100, 250, 500, 750$, and 1000 . We performed 1000 replications.

Table 1 and 2 report the sample MSE and bias (in parenthesis) of each estimator from 1000 replications:

$$\frac{1}{1000} \sum_{r=1}^{1000} \left(\hat{\omega}_{ij}^{(r)} - \int_0^T \omega_{ij}^{(r)}(t) dt \right)^2 \quad \text{and} \quad \frac{1}{1000} \sum_{r=1}^{1000} \left(\hat{\omega}_{ij}^{(r)} - \int_0^T \omega_{ij}^{(r)}(t) dt \right),$$

where r denotes the number of replications.

Figure 3, 4, and 5 show the distribution of normalized errors of each estimator:

$$\frac{\hat{\omega}_{11} - \int_0^T \omega_{11}(t) dt}{\int_0^T \omega_{11}(t) dt}, \quad \frac{\hat{\omega}_{22} - \int_0^T \omega_{22}(t) dt}{\int_0^T \omega_{22}(t) dt}, \quad \text{and}, \quad \frac{\hat{\omega}_{12} - \int_0^T \omega_{12}(t) dt}{\int_0^T \omega_{12}(t) dt},$$

respectively.

Because 1st asset is more high-frequency sampled (average duration is 45 seconds) than 2nd asset (average duration is 60 seconds), each estimate of $\int_0^T \omega_{11}(t) dt$ is more accurate than that of $\int_0^T \omega_{22}(t) dt$.

⁷Of course, our method allows the duration to be correlated or autocorrelated. See Engle and Russell (1998) for an autocorrelated duration model.

Table 1: Sample MSE from 1000 ‘daily’ replications

	$\int_0^T \omega_{11}(t) dt$	$\int_0^T \omega_{12}(t) dt$	$\int_0^T \omega_{22}(t) dt$
10-min	11.045 (0.0047)	8.9639 (-1.543)	10.829 (0.0091)
5-min	5.9340 (-0.002)	12.559 (-3.059)	6.1286 (0.0337)
2-min	3.0546 (-0.018)	48.079 (-6.836)	3.4183 (0.0091)
FE	2.3386 (0.0090)	5.8408 (-0.949)	2.6143 (0.0171)
RV	2.0397 (0.0051)	2.2274 (-0.044)	2.4936 (-0.017)
NW	2.0360 (-0.021)	2.2258 (-0.055)	2.4892 (-0.053)
OW	0.6893 (-0.045)	1.1077 (-0.103)	0.9047 (-0.085)

Note: Sample biases are given in parentheses. 10-min: $\hat{\omega}_{ij}^P(144)$; 5-min: $\hat{\omega}_{ij}^P(288)$; 2-min: $\hat{\omega}_{ij}^P(720)$; FE: $\hat{\omega}_{11}^F$ and $\hat{\omega}_{22}^F$ with $Q = 1000$, $\hat{\omega}_{12}^F$ with $Q = 100$; NE: $\hat{\omega}_{ij}^N$; OW: weighted realized volatility using optimal weights.

Under our simulation design, the correlation between the 1st and 2nd asset is on average positive: $\omega_{12}(t)$ varies around a positive mean of 0.0002 because

$$\omega_{12}(t) = \sigma_{11}(t)\sigma_{21}(t) + \sigma_{12}(t)\sigma_{22}(t)$$

and each σ_{ij} has the mean of 0.01. As expected from the bias (2.5), the shorter the interpolation time intervals is, the more downward biased the previous tick interpolation realized cross volatility $\hat{\omega}_{12}^P$ is. On the other hand, as the partitions get finer and finer, $\hat{\omega}_{11}^P(M)$ and $\hat{\omega}_{22}^P(M)$ become more accu-

Table 2: Sample MSE of Fourier estimators

Q	$\int_0^T \omega_{11}(t) dt$	$\int_0^T \omega_{12}(t) dt$	$\int_0^T \omega_{22}(t) dt$
10	71.593 (0.1165)	48.191 (0.0013)	66.515 (-0.124)
25	29.517 (0.0472)	20.286 (-0.094)	30.028 (-0.088)
50	15.984 (0.0382)	9.9519 (-0.248)	14.389 (-0.049)
100	8.2514 (-0.166)	5.8408 (-0.949)	7.7023 (-0.064)
250	3.8816 (-0.027)	14.618 (-3.580)	3.9069 (0.0488)
500	2.7194 (-0.003)	56.515 (-7.446)	3.0274 (0.0400)
750	2.3813 (-0.004)	97.244 (-9.819)	2.7696 (0.0405)
1000	2.3386 (0.0090)	128.65 (-11.31)	2.6143 (0.0171)

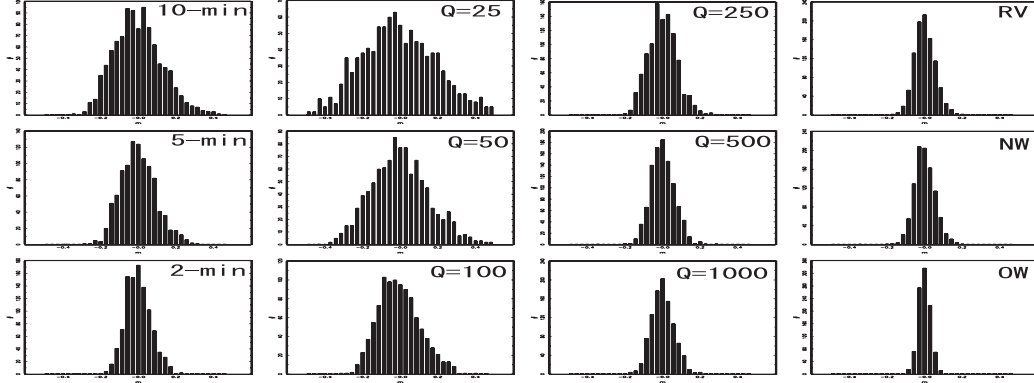
Note: Sample biases are given in parentheses.

rate. If $M \rightarrow \infty$ (in this case, $M = 60 \times 60 \times 24$), $\hat{\omega}_{11}^P(M)$ and $\hat{\omega}_{22}^P(M)$ are exactly consistent with $\hat{\omega}_{11}^R$ and $\hat{\omega}_{22}^R$, respectively.

This relationship between previous tick realized volatility and the number of partition is similar to that between Fourier series estimator and the number of Fourier coefficients. As mentioned in 2.3, as $Q \rightarrow \infty$, $\hat{\omega}_{11}^F$, $\hat{\omega}_{22}^F$, and $\hat{\omega}_{12}^F$ go to $\hat{\omega}_{11}^R$, $\hat{\omega}_{22}^R$, and 0, respectively. We cannot find the optimal Q for Fourier estimator of cross volatility unless the we know true process of volatility.

Since (2.12) is unbiased estimator of cross volatility, the sample bias is very small. As expected from the link between naively weighted realized

Figure 3: Distribution of normalized error (volatility of 1st asset)



Note: 10-min, 5-min, and 2-min denote $\hat{\omega}_{11}^P(M)$ with $M = 144, 288$, and 720 , respectively. “ $Q =$ ” signifies the Fourier estimator $\hat{\omega}_{11}^F$ with $Q = 25, 50, 100, 250, 500$, and 1000 . RV denotes the raw data realized volatility $\hat{\omega}_{11}^R$. NW denotes the naively weighted realized volatility $\hat{\omega}_{11}^N$. OW denotes the weighted realized volatility using the optimal weight. The distribution is computed with 1000 ‘daily’ replications.

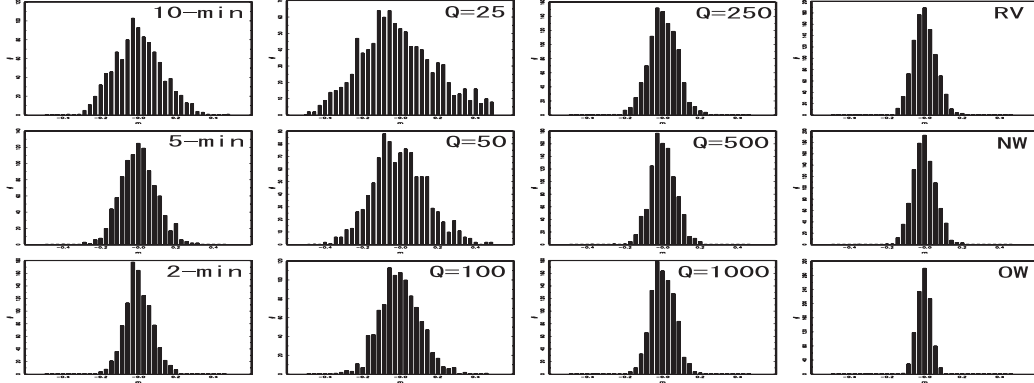
volatility $\hat{\omega}_{ij}^N$ and the raw data realized volatility $\hat{\omega}_{ij}^R$, although the former is slightly more downward biased than the latter, the former has slightly smaller sample MSE than the latter.

The optimally weighed realized volatility is overwhelming the other method. The results of optimally weighted realized volatility show principal limit of the weighted realized volatility. One of the most important remaining works is to investigate the other feasible weighting schemes by using the framework of the optimal weight.

5 Concluding remarks

In this paper we propose the definition of weighted realized volatility which nests various estimators and show some important examples. The definition is useful to make a comparative study on them. As a natural consequence,

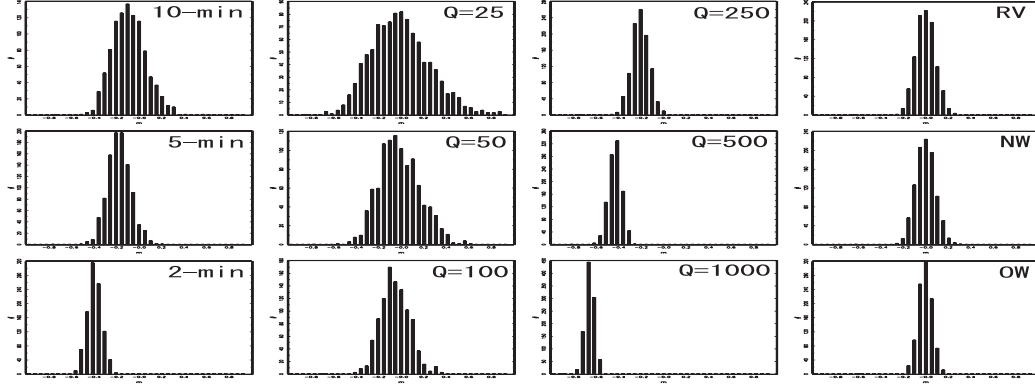
Figure 4: Distribution of normalized error (volatility of 2nd asset)



Note: 10-min, 5-min, and 2-min denote $\hat{\omega}_{22}^P(M)$ with $M = 144, 288$, and 720 , respectively. “ $Q =$ ” signifies the Fourier estimator $\hat{\omega}_{22}^F$ with $Q = 25, 50, 100, 250, 500$, and 1000 . RV denotes the raw data realized volatility $\hat{\omega}_{22}^R$. NW denotes the naively weighted realized volatility $\hat{\omega}_{22}^N$. OW denotes the weighted realized volatility using the optimal weight. The distribution is computed with 1000 ‘daily’ replications.

we derive the MSE-minimizing estimator in the class. To construct it, the estimates of optimal weights are required. We propose a feasible example of it. However, it is one of the remaining works to refine upon the feasible estimator. Another remaining work is the correction of interpolation bias. It is necessary when we can just obtain evenly spaced data which have already been interpolated.

Figure 5: Distribution of normalized error (cross volatility)



Note: 10-min, 5-min, and 2-min denote $\hat{\omega}_{12}^P(M)$ with $M = 144, 288$, and 720 , respectively. “ $Q =$ ” signifies the Fourier estimator $\hat{\omega}_{12}^F$ with $Q = 25, 50, 100, 250, 500$, and 1000 . RV denotes the raw data realized volatility $\hat{\omega}_{12}^R$. NW denotes the naively weighted realized volatility $\hat{\omega}_{12}^N$. OW denotes the weighted realized volatility using the optimal weight. The distribution is computed with 1000 ‘daily’ replications.

A Appendix

A.1 Weight matrix of $\hat{\omega}_{ii}^L$

Using

$$\alpha_m + \beta_m = 1, \quad \text{and}$$

$$p_i(t_k^i) = p_i(t_0) + \sum_{l=1}^k \Delta p_i(t_l^i)$$

we obtain

$$\begin{aligned}
& \hat{\omega}_{ii}^L(3) \\
&= \left(q_i \left(\frac{T}{3} \right) - q_i(0) \right)^2 + \left(q_i \left(\frac{2T}{3} \right) - q_i \left(\frac{T}{3} \right) \right)^2 + \left(q_i(T) - q_i \left(\frac{2T}{3} \right) \right)^2 \\
&= \{ \alpha_1 p_i(t_3^i) + \beta_1 p_i(t_2^i) - p_i(t_0) \}^2 \\
&+ \{ \alpha_2 p_i(t_7^i) + \beta_2 p_i(t_6^i) - \alpha_1 p_i(t_3^i) - \beta_1 p_i(t_2^i) \}^2 \\
&+ \{ p_i(t_8^i) - \alpha_2 p_i(t_7^i) - \beta_2 p_i(t_6^i) \}^2 \\
&= \left\{ \sum_{k=1}^2 \Delta p_i(t_k^i) + \alpha_1 \Delta p_i(t_3^i) \right\}^2 \\
&+ \left\{ \beta_1 \Delta p_i(t_3^i) + \sum_{k=3}^6 \Delta p_i(t_k^i) + \alpha_2 \Delta p_i(t_7^i) \right\}^2 \\
&+ \{ \beta_2 \Delta p_i(t_7^i) + \Delta p_i(t_8^i) \}^2. \tag{A.1}
\end{aligned}$$

Each coefficient of $\Delta p_i(t_k^i) \Delta p_i(t_l^i)$ in (A.1) is equivalent to the kl element of (2.14).

A.2 Weighted realized volatility representation of Fourier estimator

Fourier coefficients of $a_q(dp_i)$ and $b_q(dp_i)$ are approximated by

$$\begin{aligned}
a_q(dp_i) &\approx \frac{1}{\pi} \sum_{k=1}^{N_i} \cos qt_k^i \Delta p_i(t_k^i) \\
b_q(dp_i) &\approx \frac{1}{\pi} \sum_{k=1}^{N_i} \sin qt_k^i \Delta p_i(t_k^i),
\end{aligned}$$

respectively. By these approximates and the additional theorem,

$$\begin{aligned}
\hat{\omega}_{ij}^F &= \frac{\pi^2}{Q} \sum_{q=1}^Q (a_q(dp_i)a_q(dp_j) + b_q(dp_i)b_q(dp_j)) \\
&= \frac{1}{Q} \sum_{q=1}^Q \left\{ \sum_{k=1}^{N_i} \cos(qt_k^i) \Delta p_i(t_k^i) \sum_{l=1}^{N_j} \cos(qt_l^j) \Delta p_j(t_l^j) \right\} \\
&\quad + \frac{1}{Q} \sum_{q=1}^Q \left\{ \sum_{k=1}^{N_i} \sin(qt_k^i) \Delta p_i(t_k^i) \sum_{l=1}^{N_j} \sin(qt_l^j) \Delta p_j(t_l^j) \right\} \\
&= \frac{1}{Q} \sum_{q=1}^Q \left\{ \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} \{ \cos(qt_k^i) \cos(qt_l^j) + \sin(qt_k^i) \sin(qt_l^j) \} \Delta p_i(t_k^i) \Delta p_j(t_l^j) \right\} \\
&= \frac{1}{Q} \sum_{q=1}^Q \left\{ \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} \cos q(t_k^i - t_l^j) \Delta p_i(t_k^i) \Delta p_j(t_l^j) \right\} \\
&= \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} \Delta p_i(t_k^i) \Delta p_j(t_l^j) \left\{ \frac{\sum_{q=1}^Q \cos q(t_k^i - t_l^j)}{Q} \right\}.
\end{aligned}$$

Since

$$\sum_{q=1}^Q \cos qx = \begin{cases} Q & \text{if } x = 0 \\ \frac{\sin \frac{(Q+1)x}{2} \cos \frac{Qx}{2}}{\sin \frac{x}{2}} & \text{otherwise} \end{cases},$$

we get the desired result.

A.3 Variance of $\hat{\omega}_{ij}$

Using

$$\begin{aligned}
&E \left(\Delta p_i(t_k^i) \Delta p_j(t_l^j) w_{kl}^A \right)^2 \\
&= (w_{kl}^A)^2 \left\{ 2 \left(\int_{I(k,l)} \omega_{ij} dt \right)^2 + \left(\int_{t_{k-1}}^{t_k} \omega_{ii} dt \right) \left(\int_{t_{l-1}}^{t_l} \omega_{jj} dt \right) \right\},
\end{aligned}$$

$$\begin{aligned}
& V(\hat{\omega}_{ij}) \\
&= V\left(\sum_A \Delta p_i(t_k^i) \Delta p_j(t_l^j) w_{kl} + \sum_{A^C} \Delta p_i(t_k^i) \Delta p_j(t_l^j) w_{kl}\right) \\
&= E\left(\sum_A \Delta p_i(t_k^i) \Delta p_j(t_l^j) w_{kl}\right)^2 - \left\{E\left(\sum_A \Delta p_i(t_k^i) \Delta p_j(t_l^j) w_{kl}\right)\right\}^2 \\
&\quad + E\left(\sum_{A^C} \Delta p_i(t_k^i) \Delta p_j(t_l^j) w_{kl}\right)^2 \\
&= \sum (w_{kl}^A)^2 \left\{\left(\int_{I(k,l)} \omega_{ij} dt\right)^2 + \left(\int_{t_{k-1}}^{t_k} \omega_{ii} dt\right) \left(\int_{t_{l-1}}^{t_l} \omega_{jj} dt\right)\right\} \\
&\quad + \left\{\sum w_{kl}^A \int_{I(k,l)} \omega_{ij} dt\right\}^2 - \left\{\sum w_{kl}^A \int_{I(k,l)} \omega_{ij} dt\right\}^2 \\
&\quad + \sum \left(\int_{t_{k-1}}^{t_k} \omega_{ii} dt\right) \left(\int_{t_{l-1}}^{t_l} \omega_{jj} dt\right) (w_{kl}^{A^C})^2 \\
&= \sum (w_{kl}^A)^2 \left\{\left(\int_{I(k,l)} \omega_{ij} dt\right)^2 + \left(\int_{t_{k-1}}^{t_k} \omega_{ii} dt\right) \left(\int_{t_{l-1}}^{t_l} \omega_{jj} dt\right)\right\} \\
&\quad + \sum (w_{kl}^{A^C})^2 \left(\int_{t_{k-1}}^{t_k} \omega_{ii} dt\right) \left(\int_{t_{l-1}}^{t_l} \omega_{jj} dt\right).
\end{aligned}$$

A.4 Proof of Theorem 3

The first order condition is

$$\frac{\partial MSE}{\partial w} = 2Dw + 2xx'w - 2xx'1,$$

then we get

$$\begin{aligned}
w &= (D + xx')^{-1} xx'1 \\
&= \left(D^{-1} - \frac{1}{1 + x'D^{-1}x} D^{-1}xx'D^{-1} \right) xx'1 \\
&= \begin{pmatrix} \frac{\int_0^T \omega_{ij} dt \int_{I(1,1)} \omega_{ij} dt}{v_{11}\{1+\sum u_{kl}\}} \\ \vdots \\ \frac{\int_0^T \omega_{ij} dt \int_{I(k,l)} \omega_{ij} dt}{v_{kl}\{1+\sum u_{kl}\}} \\ \vdots \\ \frac{\int_0^T \omega_{ij} dt \int_{I(N_i, N_j)} \omega_{ij} dt}{v_{N_i N_j}\{1+\sum u_{kl}\}} \end{pmatrix}.
\end{aligned}$$

The second equality follows from the updating formula. See e.g., Greene (1999). The second order derivative matrix is

$$\frac{\partial^2 MSE}{\partial w \partial w'} = 2D + 2xx'.$$

This matrix is positive definite. Substituting the optimal weight into (3.1) and (3.2), we obtain (3.4) and (3.5), respectively.

References

- ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2003): “Modeling and Forecasting Realized Volatility,” *Econometrica*, 71, 579–625.
- BARNDORFF-NIELSEN, O. E., AND N. SHEPHARD (2004): “Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics,” *Econometrica*, 72, 885–925.
- BARUCCI, E., AND R. RENÒ (2002): “On Measuring Volatility of Diffusion Processes with High Frequency Data,” *Economics Letters*, 74, 371–378.

- BLACK, F., AND M. SCHOLES (1973): “The Pricing of Options and Corporate liabilities,” *Journal of Political Economy*, 81, 637–654.
- BOLLERSLEV, T. (1986): “Generalized Autoregressive Conditional Heteroskedasticity,” *Journal of Econometrics*, 31, 307–328.
- DACOROGNA, M. M., R. GENÇAY, U. MÜLLER, R. B. OLSEN, AND O. V. PICTET (2001): *An Introduction to High-Frequency Finance*. Academic Press.
- DROST, F. C., AND T. E. NIJMAN (1993): “Temporal Aggregation of GARCH Processes,” *Econometrica*, 61, 909–927.
- DROST, F. C., AND B. J. WERKER (1996): “Closing the GARCH Gaps: Continuous Time GARCH Modeling,” *Journal of Econometrics*, 74, 31–57.
- ENGLE, R. F. (1982): “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation,” *Econometrica*, 50, 987–1008.
- ENGLE, R. F., AND J. R. RUSSELL (1998): “Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data,” *Econometrica*, 66, 1127–1162.
- FLORENS-ZMIROU, D. (1993): “On Estimating the Diffusion Coefficient from Discrete Observations,” *Journal of Applied Probability*, 30, 790–804.
- GHYSELS, E., A. C. HARVEY, AND E. RENAULT (1996): “Stochastic Volatility,” in *Handbook of Statistics*, ed. by G. S. Maddala, and C. R. Rao, vol. 14. North Holland, Amsterdam.
- GREENE, W. H. (1999): *Econometric Analysis: Forth Edition*. Prentice Hall.

- HAYASHI, T., AND N. YOSHIDA (2005): “On Covariance Estimation of Non-Synchronously Observed Diffusion Processes,” *Bernoulli*, 11, 359–379.
- HULL, J., AND A. WHITE (1987): “The Pricing of Options on Assets with Stochastic Volatilities,” *Journal of Finance*, 42, 281–300.
- KANATANI, T. (2004): “Integrated Volatility Measuring from Unevenly Sampled Observations,” *Economics Bulletin*, 3(36), 1–8.
- LO, A. W. (1988): “Maximum Likelihood Estimation of Generalized Ito Process with Discretely Sampled Data,” *Econometric Theory*, 4, 231–247.
- MALLIAVIN, P., AND M. E. MANCINO (2002): “Fourier Series Method for Measurement of Multivariate Volatilities,” *Finance and Stochastics*, 6, 49–61.
- NELSON, D. B. (1991): “Conditional Heteroskedasticity in Asset Returns: A New Approach,” *Econometrica*, 59, 347–370.
- POON, S.-H., AND C. W. J. GRANGER (2003): “Forecasting Volatility in Financial Markets: A Review,” *Journal of Economic Literature*, 41, 478–539.
- SUEISHI, N. (2004): “Estimation of Discretely Sampled Diffusion Processes,” In *Proceedings of the 11th Kansai Econometrics Workshop*, Kobe University, January 10, 2004, 1, 5–27.